

Roßbach, Hans-Günther

## Unterrichtsqualität im 2. Schuljahr - Ergebnisse einer empirischen Untersuchung

*Unterrichtswissenschaft 30 (2002) 3, S. 230-245*



Quellenangabe/ Reference:

Roßbach, Hans-Günther: Unterrichtsqualität im 2. Schuljahr - Ergebnisse einer empirischen

Untersuchung - In: Unterrichtswissenschaft 30 (2002) 3, S. 230-245 - URN: urn:nbn:de:0111-opus-76873

- DOI: 10.25656/01:7687

<https://nbn-resolving.org/urn:nbn:de:0111-opus-76873>

<https://doi.org/10.25656/01:7687>

in Kooperation mit / in cooperation with:

# BELTZ JUVENTA

<http://www.juventa.de>

### Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.  
This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

### Kontakt / Contact:

peDOCS  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Digitalisiert

Mitglied der

  
Leibniz-Gemeinschaft

---

# Unterrichtswissenschaft

Zeitschrift für Lernforschung  
30. Jahrgang / 2002 / Heft 3

---

## Thema: Unterrichtsqualität

70 + 1.100

Verantwortlicher Herausgeber:  
Wolfgang Einsiedler

Wolfgang Einsiedler: Editorial: Das Konzept Unterrichtsqualität	194
Hartmut Ditton: Unterrichtsqualität – Konzeptionen, methodische Überlegungen und Perspektiven	197
Ferdinand Eder: Unterrichtsklima und Unterrichtsqualität	213
Hans-Günther Roßbach: Unterrichtsqualität im 2. Schuljahr – Ergebnisse einer empirischen Untersuchung	230
Marten Clausen, Kai Schnabel, Sabine Schröder: Konstrukte der Unterrichtsqualität im Expertenurteil	246
Andreas Helmke: Kommentar: Unterrichtsqualität und Unterrichtsklima – Perspektiven und Sackgassen	261

## Allgemeiner Teil

Jens Möller: Informationsbedingte Veränderungen elterlicher Einstellungen zu Berichtszeugnissen: eine experimentelle Studie	278
---	-----

---

Hans-Günther Roßbach

## **Unterrichtsqualität im 2. Schuljahr - Ergebnisse einer empirischen Untersuchung**

Quality of Instruction in Second Grade Classrooms –  
Results of an Empirical Study

---

*Der Beitrag berichtet über Ergebnisse einer Untersuchung der Unterrichtsqualität in zweiten Grundschulklassen. Die Stichprobe besteht aus 214 Grundschulklassen und 272 8-jährigen Schülern. Im Mittelpunkt stehen Beobachtungen der Unterrichtsqualität durch externe Beobachter. Der Beitrag beschreibt das Qualitätskonzept der Studie und berichtet Ergebnisse zur faktoriellen Struktur der erhobenen Merkmale der Unterrichtsqualität, zur messtechnischen Qualität der Erhebungsinstrumente, zur Verteilung der Merkmale der Unterrichtsqualität, zu ihren Abhängigkeiten von strukturellen Rahmenbedingungen der Klasse und den pädagogischen Orientierungen der Klassenlehrerinnen sowie zu den Auswirkungen der Unterrichtsqualität auf die Schulleistungen der Schüler im Alter von 8 Jahren.*

*The paper reports on results of a study of the quality of instruction in second grade classrooms. The sample consists of 214 primary classrooms and 272 8-year-old pupils. The focus of interest are ratings of the quality of instruction by external raters. The paper describes the concept of quality and reports on results with regard to the factor structure of the measured characteristics of quality of instruction, the technical quality of the used observational instruments, the distribution of the quality characteristics, their dependencies on structural characteristics of the classroom and on educational orientations of the class teacher and on the effects of the quality of instruction on school achievement of the pupils at age 8.*

### **1. Einleitung**

Empirische Untersuchungen haben eine Vielzahl von Bedingungen identifiziert, die für erfolgreiche schulische Lernprozesse von Schülern bedeutsam sind (vgl. Helmke & Weinert, 1997a; Wang, Haertel & Walberg, 1993). Diese Bedingungen beziehen sich auf verschiedene Ebenen, von Schülervoraussetzungen und familialen Merkmalen über Unterrichts-, Klassen- und Schulmerkmale bis hin zu Kontextmerkmalen von Schulen. Auf der institutionellen Seite sind die Merkmale auf der Ebene der Klasse und des Unterrichts von zentraler Bedeutung, da sie gewissermaßen am „nächsten“ an den Lernprozessen der Schüler liegen (vgl. Sammons, Hillman & Mortimore, 1995; Scheerens, 1992). Allerdings darf dabei die Bedeutung des Klassenkontextes nicht vernachlässigt werden, der förderliche oder hinderliche Rahmenbedingungen für die auf der Klassenebene stattfindenden Prozesse bereitstellt. Als zentrales Merkmal auf der Klassenebene wird der Unterricht und seine

Qualität betrachtet. Unterschiedliche Forschungstraditionen - verbunden mit unterschiedlichen methodischen Zugriffen - haben dabei die aktuelle Forschung zur Unterrichtsqualität beeinflusst, vor allem die Erziehungsstil- und Unterrichtsstilforschung, Untersuchungen zu Merkmalen effektiven Lehrerverhaltens, die Lehrmethodenforschung, Forschungen zur Klassenführung und Forschungen zur pädagogischen Interaktion (vgl. Clausen, 2000; Einsiedler, 1997, 2000; Gruehn, 2000). Obwohl Unterrichtsqualität zu einem eigenen Forschungsbereich geworden ist, gibt es gegenwärtig nur wenige theoretische Überlegungen und begriffliche Präzisierungen. Nach Einsiedler (2000) leidet die Konzeptualisierung von Unterrichtsqualität unter theoretischen Mängeln. Die im Einzelnen im Kontext der Unterrichtsqualität betrachteten Konstrukte lassen sich unterschiedlichen Unterrichtsmodellen zuordnen und sind auch unterschiedlich empirisch abgesichert. Gruehn (2000) z.B. setzt in ihrer Untersuchung (hier bei der Analyse von Schülerwahrnehmungen) Skalen ein, die sich den Konstrukten/Modellen Unterrichtsführung, Unterrichtstempo, Klarheit des Unterrichts, remediale Unterrichtsmaßnahmen, adaptiver Unterricht, schülerorientierter Unterricht, konstruktivistischer Unterricht und affektive Qualität der Lehrer-Schüler-Beziehung zuordnen lassen. Helmke und Weinert (1997b) unterscheiden in der SCHOLASTIK-Studie 6 Variablen der Unterrichtsqualität: Klassenführung, Strukturiertheit, Unterstützung, Förderungsorientierung, Soziales Klima und Variabilität. Die verschiedenen Unterrichtskonzepte bzw. -formen können aber nicht als sich ausschließende Alternativen betrachtet werden. Weinert (1998) weist in diesem Zusammenhang auf die Notwendigkeit einer Kombination verschiedener Unterrichtsformen hin, um multiple Lernziele zu erreichen.

Vor diesem Hintergrund berichtet der folgende Beitrag über Ergebnisse einer Untersuchung der Unterrichtsqualität in deutschen Grundschulklassen (überwiegend 2. Klassen). Über die tatsächliche Ausgestaltung der Unterrichtsqualität in der Grundschule liegen nur wenige empirische Informationen vor (vgl. Einsiedler, 1997; Helmke & Weinert, 1997b); dementsprechend hat die vorliegende Untersuchung eher einen heuristischen Charakter. Die Untersuchung ist Teil der längsschnittlich angelegten European Child Care and Education (ECCE) Study, in der Kinder vom 4. bis zum 8. Lebensjahr in ihren Umwelten Familie, Kindergarten und Grundschule verfolgt werden (ECCE-Study Group 1999; Tietze, Meischner, Gänßfuß, Grenner, Schuster, Völkel & Roßbach, 1998).<sup>1</sup> Die Darstellung ist auf den deutschen Teil der Studie und die Grundschulphase beschränkt.

---

<sup>1</sup> Der deutsche Teil der European Child Care and Education Study wurde durch Sachbeihilfen des Bundesministeriums für Bildung, Wissenschaft, Forschung und Technologie und teilweise durch die EU unterstützt. Die Datenerhebungen in den Grundschulklassen wurden durch eine Sachbeihilfe der DFG ermöglicht.

## 2. Qualitätskonzept der Untersuchung

Das Qualitätskonzept unterscheidet auf der Ebene einer Grundschulklasse 3 Bereiche der „pädagogischen Qualität“:<sup>2</sup>

- *Prozessqualität* als die Qualität der in der Klasse stattfindenden pädagogischen Prozesse (*Unterrichtsqualität* im Engeren, z.B. Strukturiertheit des Unterrichts, Unterstützung der Schüler, Klassenmanagement, soziales Klima)
- *Strukturqualität* als die Qualität der Rahmenbedingungen der Klasse (z.B. Alter, Erfahrungen und wöchentliche Arbeitszeiten der Klassenlehrerinnen; Klassengröße, Anteil ausländischer Schüler; Unterrichtsstunden pro Woche; materielle und räumliche Ausstattung)
- *Qualität pädagogischer Orientierungen* als die pädagogischen Werte, Überzeugungen und Einstellungen der Klassenlehrerinnen (z.B. Vorstellungen der Lehrerinnen über die Entwicklung der Schüler, leistungsbezogene Einstellungen der Lehrerinnen)

Der Bereich der Prozess-/Unterrichtsqualität wird insofern als Zentralbereich pädagogischer Qualität angesehen, als er unmittelbar die Auseinandersetzungen der Schüler mit den jeweiligen Lernaufgaben beeinflusst. Hypothetisch wird für den Grundschulunterricht davon ausgegangen, dass die Prozessmerkmale dann das Lernen der Schüler positiv befördern, wenn sie zu Lernprozessen führen, in denen

- die Schüler aktiv die ihnen zur Verfügung stehende Lernzeit nutzen,
- sie an sachlich angemessenen Lernaufgaben arbeiten können, die sich an ihr Vorwissen anschließen und deren Vernetzungen zu vorherigen, nachfolgenden und verwandten Inhalten ihnen deutlich ist,
- die Schüler die sozial-emotionale Atmosphäre als angenehm, zumindest nicht als bedrohend wahrnehmen und sie sich als Person angenommen fühlen,
- die Anwendungskontexte der Lernaufgaben den Schüler deutlich sind bzw. sie sie in verschiedenen, ihnen einsichtigen, auch authentischen Situationen üben können und
- die Schüler ihre Lernprozesse reflektieren können und sie kognitive und metakognitive Lernstrategien auf verschiedenen Ebenen erfahren.

---

<sup>2</sup> Die gleiche Konzeptualisierung von pädagogischer Qualität (jeweils mit setting-spezifischen Messinstrumenten operationalisiert) wurde in der ECCE-Studie auch für alle analysierten Umwelten benutzt, die die Untersuchungskinder im Längsschnitt erfahren haben: Familie und Kindergarten (im Alter von 4 Jahren) sowie Familie und Grundschulklasse (im Alter von 8 Jahren).

### 3. Untersuchungsdesign und Fragestellungen

In über 80% der Fälle befanden sich die untersuchten Kinder im Alter von 8 Jahren in der 2. Grundschulklasse. In einer etwa einstündigen standardisierten mündlichen Befragung der Klassenlehrerinnen wurden die Rahmenbedingungen der Klasse sowie die pädagogischen Orientierungen der Lehrerinnen erfasst. Das Befragungsinstrument wurde speziell für diese Untersuchung konstruiert und besteht weitgehend aus Fragen mit geschlossenen Antwortvorgaben. Die Prozessmerkmale wurden über Beobachtungen durch externe und im Rahmen des Projekts ausführlich trainierte Beobachter erfasst. Lehrer- und Schülerwahrnehmungen der Prozesse konnten nicht erhoben werden. Die Beobachtungsverfahren konnten für die Untersuchung nicht spezifisch konstruiert werden, vielmehr musste auf bereits bestehende Instrumente zurückgegriffen werden. Ausgewählt wurden 2 hochinferente Verfahren:

- *Unterrichtsbeurteilungen durch externe Beobachter* - im Folgenden abgekürzt: UEB (Helmke & Schrader, 1997): Eingesetzt wurden 25 der insgesamt 27 fünfstufigen Skalen dieses Ratingverfahrens (umgepolt gegenüber Helmke & Schrader; hier von 1 = „trifft gar nicht zu“ bis 5 = „trifft vollkommen zu“). Helmke und Schrader haben 11 der Items zu 6 Subskalen gruppiert: Klassenführung, Strukturierung, Unterstützung, Förderungsorientierung, Soziales Klima und Variabilität. Für die vorliegende Untersuchung wurden für jedes Item itemspezifische Beispiele hinzugefügt.
- *Instructional Environment Observation Scales - IEOS* (Secada, 1997): Dieses Beobachtungsinstrument besteht aus verschiedenen Ratingskalen, die ebenfalls fünfstufig aufgebaut sind, wobei jede Skalenstufe ausführlich beschrieben wird. Die IEOS orientieren sich an einem „kindorientierten“ Unterrichtsmodell und an dem Konzept der „authentischen Instruktion“ und betonen damit - als pragmatischer Grundkonsens der verschiedenen konstruktivistischen Ansätze (vgl. Gerstenmaier & Mandl, 1995) - die stärkere Einbeziehung bedeutungsvoller Kontexte und authentischer Aufgaben, speziell inwieweit den Schülern die Bezüge eines behandelten Gegenstandes zu den Anwendungssituationen außerhalb der Schule sowie fächerübergreifende Perspektiven verdeutlicht werden. Eingesetzt wurden 6 Skalen: Klassenklima, Klassenroutinen, fächerübergreifende Verbindungen, Verbindung zum Leben außerhalb der Klasse, soziale Unterstützung für das Lernen der Schüler, Schülerengagement.

Die Beobachtungsskalen werden von den jeweiligen Autoren als intervallskaliert betrachtet. Beide Beobachtungsverfahren beziehen sich nur auf allgemeine Aspekte des Unterrichts, d.h., sie beinhalten keine fachspezifischen oder fachdidaktischen Teile. Das gesamte System der IEOS enthält zwar fachspezifische Teile für Unterricht in Lesen und Schreiben sowie Unterricht in Mathematik; erste Erprobungen der zum Erhebungszeitpunkt zur Ver-

fügung stehenden Version zeigten aber erhebliche Bodeneffekte der entsprechenden Items auf, sodass - in internationaler Absprache - auf den Einsatz dieser fachspezifischen Teile verzichtet werden musste. Beide Instrumente wurden am Ende der Beobachtung von 2 Unterrichtsstunden an einem Vormittag durch trainierte Beobachter bearbeitet. Der beobachtete Unterricht musste von der Klassenlehrerin erteilt werden und - wenn möglich, was aber in der Regel gelang - Unterricht in Mathematik *und* Lesen/Schreiben enthalten. Auf eine fachspezifische Erfassung der Prozessqualität musste verzichtet werden. Aus untersuchungsökonomischen Gründen war für die Gesamtstichprobe der Klassen nur eine einmalige Erhebung der Prozessqualität möglich, für eine kleine Subgruppe (vgl. weiter unten) wurde der Unterricht zweimal beobachtet.

In der ECCE-Studie wird der Entwicklungsstand der Kinder zu verschiedenen Zeitpunkten breit erfasst, und zwar sowohl im Hinblick auf die kognitive als auch auf die sozial-emotionale Entwicklung. Der vorliegende Beitrag beschränkt sich auf den Ausschnitt der Schulleistungen im Alter von 8 Jahren.<sup>3</sup> Die Schulleistungen wurden durch eine deutsche Adaptation des Woodcock Johnson Test of Achievement WJ (Woodcock & Johnson, 1989, 1990) gemessen; ausgewählt wurden die Subtests Satzverständnis, Rechnen, Lösen angewandter Probleme und Kenntnisse über die natürliche und soziale Umwelt. Zu Beginn des Längsschnitts bestand die Stichprobe aus 422 Kindergartenkindern im Alter von 4 Jahren aus 103 Kindergartengruppen (in der Regel 4 Kinder pro Gruppe) aus 103 Kindergärten in Baden-Württemberg, Berlin Ost und West, Brandenburg, Nordrhein-Westfalen und Rheinland-Pfalz. Im Alter von 8 Jahren konnten Daten von 306 (72,5%) dieser Kinder erhoben werden. In diesem Beitrag werden nur die Kinder berücksichtigt, für die vollständige Unterrichtsbeobachtungen vorliegen (272 Kinder). Die Kinder verteilen sich auf 214 Grundschulklassen in den genannten Bundesländern. Aufgrund dieses Längsschnittdesigns liegen somit nur wenige Schülerdaten pro Klasse vor, in etwas mehr als 50% der Klassen nur die Daten von einem Schüler.

Folgende allgemeine Fragestellungen stehen im Mittelpunkt des Beitrags: Wie ist die messtechnische Qualität und die faktorielle Struktur der eingesetzten Instrumente zur Erfassung der Unterrichtsqualität in der Grundschule? Wie verteilen sich die verschiedenen Aspekte der Unterrichtsqualität? Inwieweit hängt die Unterrichtsqualität von strukturellen Rahmenbedingungen und den Orientierungen der Lehrkräfte ab? Welche Beziehungen der Unterrichtsqualität zu den Schulleistungen im Alter von 8 Jahren gibt es?

---

<sup>3</sup> Die anderen Entwicklungsbereiche werden in weiteren Veröffentlichungen analysiert werden.

## 4. Ergebnisse

### 4.1 Faktorielle Struktur und technische Qualität der eingesetzten Skalen zur Unterrichtsqualität

Zur Aufklärung der Dimensionierung der eingesetzten Verfahren zur Erfassung der Unterrichtsqualität (Prozessqualität) wurden auf der Ebene der 214 Grundschulklassen die insgesamt 31 Items der beiden Beobachtungsverfahren (25 UEB und 6 IEOS) gemeinsamen Faktorenanalysen unterzogen. Die verschiedenen Kriterien zur Bestimmung der Faktorenanzahl weisen auf 4 bis 6 Faktoren hin; eine gut interpretierbare Lösung ergibt sich bei 5 Faktoren, denen sich 26 Items klar zuordnen lassen (Ladungen  $\geq |.5|$ ; im Folgenden jeweils in Klammern die Anzahl der Items mit entsprechenden Ladungen):

- *Klassenmanagement* (9 Items). Der positive Pol spricht eine Organisation des Unterrichts an, bei der Übergänge fließend sind, klare Regeln über das Verhalten in der Klasse eingehalten werden, das Schülerengagement hoch ist und effektiv mit Störungen umgegangen wird. Dazu gehört auch, dass die Lehrkraft sich um ein Klima wechselseitigen Respekts und sozialer Unterstützung des Lernens bemüht und schwächere Schüler nicht entmutigt werden.
- *Schülerorientiertes sozial-emotionales Klima* (6 Items). Der positive Pol indiziert ein Klima, in dem Schüler als Personen anerkannt werden und auch Raum für ihre privaten Gefühle und Probleme besteht. Dies geht einher mit einer nicht nur auf engere fachliche Ziele bezogenen Orientierung des Unterrichts und der Ablehnung eines leistungsorientierten und kontrollierenden Klimas.
- *Direktes Unterrichten* (6 Items). Der positive Pol indiziert einen Unterricht mit einer klaren Strukturierung, Überblicken über den Unterrichtsverlauf, Einführungen und Zusammenfassungen sowie Hervorhebungen der wichtigsten Ziel- und Inhaltsaspekte; es werden Lernstrategien vermittelt, und das Unterrichtstempo ist eher hoch.
- *Anwendungsbezug* (2 Items). Bei positiven Werten wird im Unterricht Wert darauf gelegt, bei den behandelten Inhalten, Aktivitäten oder Lernaufgaben ihre fächerübergreifenden Vernetzungen und ihre Verbindungen zu Inhalten oder Kompetenzen außerhalb des Klassenzimmers im „wirklichen Leben“ aufzuzeigen.
- *Unterstützung bei Stillarbeit* (3 Items). Dieser Faktor bezieht sich auf einen speziellen Aspekt des Unterrichts. Der positive Pol indiziert eine aktive Unterstützung der Schüler durch die Lehrkraft in Stillarbeitsphasen.

Die entsprechend der Markiertvariablen additiv gebildeten Skalen sind weitgehend unabhängig voneinander (Interkorrelationen von  $-.09$  bis  $.35$ ). Die internen Konsistenzen fallen in Anbetracht der teilweise nur wenigen Items,



die in eine Skala eingehen, befriedigend aus (Cronbachs Alpha für Klassenmanagement .92, Schülerorientiertes sozial-emotionales Klima .77, Direktes Unterrichten .73, Anwendungsbezug .50, Unterstützung bei Stillarbeit .70; für einen aus beiden Instrumenten gebildeten Gesamtwert .87).

Eine einmalige zweistündige Beobachtung der Unterrichtsqualität kann möglicherweise sehr fehleranfällig sein. Zur Abschätzung der möglichen Fehler wurden in einem Teil der Klassen die Beobachtungen zweimal durchgeführt: In 21 Klassen konnten zum gleichen Zeitpunkt 2 Beobachter eingesetzt werden (Beobachterübereinstimmung). In 20 Klassen hat der gleiche Beobachter zu 2 etwa 3 bis 4 Wochen auseinanderliegenden Zeitpunkten die Unterrichtsqualität eingeschätzt (Stabilität, gleicher Beobachter), und für 21 Klassen liegen die Einschätzungen von verschiedenen Beobachtern zu 2 ebenfalls etwa 3 bis 4 Wochen auseinanderliegenden Zeitpunkten vor (Stabilität, verschiedene Beobachter). Für jedes Beobachtungspaar wurde über alle 31 Items ausgezählt, bei wie vielen Items sich eine exakte Übereinstimmung und eine Übereinstimmung bei maximal einer Skalenstufe Abweichung zeigt; anschließend wurden für jedes Paar die Übereinstimmungen in prozentuale Übereinstimmungen umgerechnet. Die exakten Beobachterübereinstimmungen reichen von einer Klasse mit 23% (exakte Übereinstimmung in 7 der 31 Items) bis zu einer Klasse mit 94% (exakte Übereinstimmung in 29 Items); der Median der Beobachterübereinstimmung über die 21 Klassen liegt bei 42%. Betrachtet man bei fünfstufigen Skalen eine Abweichung von maximal einem Skalenpunkt noch als - zu akzeptierende - Übereinstimmung, so steigt der Median auf 94% (Spannweite von 68% bis 100%). Insgesamt kann deshalb von einer guten Beobachterübereinstimmung ausgegangen werden.

Eine hohe Übereinstimmung der Einschätzungen zu 2 Zeitpunkten (gleicher oder verschiedene Beobachter) weist sowohl auf eine hohe Reliabilität des Beobachtungsverfahrens als auch auf eine hohe Stabilität des Merkmals selbst hin. Zu berücksichtigen ist im vorliegenden Fall allerdings, dass zwischen den beiden Beobachtungszeitpunkten nur eine Zeitspanne von 3 bis 4 Wochen liegt und damit höhere Stabilitäten begünstigt werden. Bei Einsatz der gleichen Beobachter zu den 2 Zeitpunkten reichen die exakten Übereinstimmungen von 26% bis 68% (Median 47%), bei Übereinstimmung innerhalb von maximal einer Skalenstufe Abstand von 71% bis 100% (Median 87%). Die entsprechenden Werte bei Einsatz von verschiedenen Beobachtern liegen leicht niedriger (exakte Übereinstimmung von 16% bis 58%, Median 32%; Übereinstimmung innerhalb von maximal einer Skalenstufe Abstand von 58% bis 97%, Median 77%), was vermutlich auf Erinnerungseffekte bzw. mögliche Einschätzungstendenzen bei Einsatz der gleichen Beobachter zurückgeführt werden kann. Insgesamt zeigen sich über 2 verschiedene Zeitpunkte hohe Übereinstimmungen, die auf eine zufrieden stellende Reliabilität wie auch auf eine hohe Stabilität der beobachteten Merkmale der Unterrichtsqualität verweisen.

## 4.2 Verteilung der Unterrichtsqualität

Über die Verteilung der Merkmale der Unterrichtsqualität in deutschen Grundschulklassen liegen mit Ausnahme der Ergebnisse der SCHOLASTIK-Studie (bayerische 3. und 4. Klassen; Helmke & Schrader, 1997) keine Informationen vor. Tabelle 1 enthält aus der vorliegenden Untersuchung der 214 Grundschulklassen (in der Regel 2. Klassen) Mittelwerte und Standardabweichungen für die fünf Skalen sowie für den aus allen Items von UEB und IEOS gebildeten Gesamtwert (jeweils Durchschnitte über die in eine Skala eingegangenen Itemwerte). Ergänzend sind die entsprechenden Werte getrennt für Grundschulklassen in den alten und den neuen Bundesländern wiedergegeben.

Tabelle 1:  
Prozessmerkmale des Unterrichts in der Grundschulklasse;  
Mittelwerte und Standardabweichungen

		Insgesamt	Alte Länder	Neue Länder	T-Test
Klassenmanagement	M	3,6	3,6	3,8	p<.05
	SD	0,7	0,7	0,7	
Schülerorientiertes sozial-emotionales Klima	M	2,3	2,4	1,9	p<.001
	SD	0,7	0,7	0,7	
Direktes Unterrichten	M	2,6	2,6	2,7	n.s.
	SD	0,7	0,6	0,8	
Anwendungsbezug	M	1,9	2,0	1,6	p<.001
	SD	0,9	0,9	0,7	
Unterstützung bei Stillarbeit	M	3,5	3,5	3,5	n.s.
	SD	0,9	0,9	0,8	
Gesamtwert	M	3,0	3,0	3,0	n.s.
	SD	0,4	0,4	0,5	

Die Beobachtungsskalen selbst enthalten keine direkten Normen/Standards in dem Sinne, dass bei bestimmten Ausprägungen eine unzureichende, mittelmäßige oder gute Qualität indiziert ist. Aufgrund der Formulierungen der Skalen (vgl. Helmke & Schrader, 1997; Secada, 1997) kann aber vorsichtig die theoretische Mitte der fünfstufigen Skalen (3,0) als Bezugspunkt interpretiert werden, unter dem von einer eher schlechteren und über dem von einer eher besseren Qualität gesprochen werden kann. Der Mittelwert des Gesamtwerts der Unterrichtsqualität liegt mit 3,0 genau in dieser theoretischen Mitte. Dabei gibt es keine Unterschiede zwischen Klassen aus den alten und den neuen Bundesländern. Wenn sich Unterschiede zwischen Klassen in den alten und den neuen Bundesländern zeigen, so liegen diese ausschließlich bei den Subskalen. Signifikante und stärkere Unterschiede zwischen den Klassen in den alten und neuen Bundesländern ergeben sich dabei nur für das schülerorientierte sozial-emotionale Klima und den Anwendungsbezug, in denen die Unterschiede sich auf fast eine halbe Skalenstufe belaufen. Beide

Male werden diese Aspekte in den neuen Bundesländern als schlechter realisiert eingeschätzt.

Auf der Ebene der 5 Subskalen ergibt sich insgesamt und für alte und neue Bundesländer das gleiche Muster. Das Klassenmanagement und die Unterstützung bei der Stillarbeit liegen über der theoretischen Skalenmitte von 3,0; beide Aspekte des Unterrichts werden somit im Durchschnitt in den beobachteten Klassen als eher positiv realisiert eingeschätzt. Angesichts der gegenwärtigen Diskussion um die Gestaltung des Grundschulunterrichts, in der offene Lernformen deutlich präferiert werden, überrascht nicht, dass der Mittelwert für die Skala Direktes Unterrichten im unteren Bereich liegt. Überraschend ist aber - angesichts der Wertschätzung dieses Aspektes in der pädagogischen Literatur - der noch niedrigere Mittelwert im schülerorientierten sozial-emotionalen Klima, das damit in den beobachteten Klassen als eher negativ eingeschätzt wird. Der niedrigste Mittelwert ergibt sich für den Anwendungsbezug, d.h., dass im Durchschnitt weniger Wert darauf gelegt wird, bei den behandelten Inhalten, Aktivitäten oder Lernaufgaben ihre fächerübergreifenden Vernetzungen und ihre Verbindungen zu Inhalten oder Kompetenzen außerhalb des Klassenzimmers aufzuzeigen. Angesichts der theoretischen Erwartungen, dass durch die Verdeutlichung der Anwendungsbezüge die Lernmotivationen der Schüler gesteigert wie auch der Gebrauch des gelernten Wissens verbessert werden kann, weist dieses Ergebnis auf eine im Durchschnitt unbefriedigende Situation hin. Die genannten Mittelwerte verbergen aber die Unterschiede zwischen den einzelnen Klassen. Immerhin liegen die Streuungen in den Subskalen bei 0,7 bis 0,9 Skalenstufen.<sup>4</sup>

---

<sup>4</sup> Bei einem Vergleich der Ergebnisse mit denen der SCHOLASTIK-Studie (Helmeke & Schrader, 1997) zeigt sich folgendes Bild (ohne Tabelle): Bei den Items, die den Subskalen Klassenmanagement und Unterstützung bei Stillarbeit zugeordnet werden können, zeigen sich praktisch keine Unterschiede zwischen den hier untersuchten 2. Klassen und den in der SCHOLASTIK-Studie in Bayern untersuchten 3. und 4. Klassen. Im Hinblick auf Aspekte des direkten Unterrichts liegen die bayerischen Klassen aber deutlich höher (etwa um eine ganze Skalenstufe). Ebenfalls deutlich positiver werden Aspekte des schülerorientierten sozial-emotionalen Klimas eingeschätzt (rund 3/4 Skalenstufe). Auch im Hinblick auf den Einzelaspekt der Individualisierung (hier nicht bei der Bildung von Subskalen berücksichtigt) liegt die Unterrichtsqualität in den bayerischen Klassen deutlich höher. Es kann an dieser Stelle nicht beantwortet werden, ob diese Unterschiede auf unterschiedliche Beobachtertrainings oder den unterschiedlichen Gebrauch des Beobachtungsinstruments (UEB), auf die unterschiedlichen Klassenstufen oder auf eine bessere Unterrichtsqualität in den entsprechenden Aspekten in bayerischen Klassen gegenüber den hier untersuchten Klassen zurückzuführen sind.

### 4.3 Beziehungen zwischen Orientierungs-, Struktur- und Prozessqualität

Das Konzept der Qualität unterscheidet zwischen Orientierungs-, Struktur- und Prozess-/Unterrichtsqualität. Damit stellt sich die Frage, wie diese drei Bereiche zusammenhängen, speziell inwieweit die Unterrichtsqualität von den Rahmenbedingungen der Klasse und den mehr übergreifenden Orientierungen der Klassenlehrerinnen abhängt. In bisherigen Untersuchungen wurden eher niedrige Beziehungen zwischen strukturellen Merkmalen von Grundschulklassen und der Unterrichtsqualität gefunden. Helmke und Wehnert (1997b, 246f.) fanden z.B., dass Merkmale der Klassenzusammensetzung insgesamt gesehen nur eine marginale Rolle für Unterschiede in der Unterrichtsqualität spielen. Da die vorliegende Untersuchung nicht speziell auf die Frage der Abhängigkeiten der Prozessqualität angelegt war, müssen die folgenden Analysen eher als heuristische Versuche betrachtet werden. Zur Beantwortung der Frage wurde ein regressionsanalytischer Ansatz ausgewählt mit Strukturmerkmalen und pädagogischen Orientierungen der Klassenlehrerinnen als Prädiktoren und der Unterrichtsqualität als Kriterium (5 Subskalen und Gesamtwert). Ausgewählt wurde ein Satz von 19 Prädiktoren. Tabelle 2 enthält die Anteile aufgeklärter Varianzen sowie die Regressions-Faktor-Struktur-Koeffizienten (RFSK; einfache Korrelationen der Prädiktoren

Tabelle 2:  
Regressionen der Prozessqualität auf Strukturmerkmale und pädagogische Orientierungen der Klassenlehrerinnen ( $R^2$ ; nur RFSK  $\geq |.03|$ )

	Klassen- manage- ment	Sozial- emotio- nales Klima	Anwen- dungs- bezug	Gesamt- wert
<b>Pädagogische Orientierungen der Klassenlehrerinnen</b>				
Entwicklungserwartungen	-	-.40	-	-.32
Leistungsbezogene Einstellungen	-	-.62	-	-
Kreativitätsbezogene Einstellungen	.34	.42	.39	.52
<b>Strukturmerkmale</b>				
Unterrichtsverpflichtung der Klassenlehrerinnen	-	.32	-	.44
Zufriedenheit der Klassenlehrerinnen mit der Arbeit <sup>1</sup>	.46	-	-	-
Anteil ausländischer Schüler mit Sprachproblemen	-.37	-	-	-
Anzahl Lehrkräfte in der Klasse	-	-	-.38	-
Differenz zwischen jüngstem und ältestem Schüler	-	-	.32	-
Unterrichtsstunden der Klasse laut Stundenplan	-	.40	-	-
Hausaufgabenumfang pro Woche	-	-	.38	-
allg. Anteil Unterricht in Sozialform "ganze Klasse"	-	-.30	-	-
Häufigkeit schriftlicher Kontrollen <sup>2</sup>	.50	-	-	-
Mittagessen in der Schule <sup>3</sup>	-.42	.41	.48	-
Anzahl verschiedener Materialien in der Klasse	.34	.32	-	.47
Anteil aufgeklärter Varianz $R^2$	19,5%**	20,5%**	20,0%**	15,4%*

<sup>1</sup> 1 = unzufrieden, 4 = zufrieden

<sup>2</sup> 1 = nie, 6 = mehr als einmal pro Woche

<sup>3</sup> 1 = ja, 2 = nein

\*  $p < .05$ ; \*\*  $p < .01$

mit den Vorhersagevariablen). Berücksichtigt sind nur die Skalen mit signifikanten  $R^2$  und die der 19 Prädiktoren, die für eine der Skalen einen RFSK von  $\geq |.30|$  haben.

Der Prädiktorensatz erklärt mit 15% nur einen kleineren Teil der Varianz in dem Gesamtwert der Unterrichtsqualität. Bei den Subskalen Klassenmanagement, Schülerorientiertes sozial-emotionales Klima und Anwendungsbezug steigt der Anteil erklärter Varianz auf 20%; keine signifikante Varianzaufklärung ergibt sich bei den Skalen Direktes Unterrichten und Unterstützung bei Stillarbeit. Strukturmerkmale und Orientierungen der Klassenlehrerinnen determinieren somit in keiner Weise die tatsächlich zu beobachtende Unterrichtsqualität, d.h., das gleiche Unterrichtsverhalten und die gleichen Interaktionen der Lehrerinnen mit den Schülern können unter verschiedenen Rahmenbedingungen beobachtet werden, wie auch unter gleichen Rahmenbedingungen unterschiedliches Prozessgeschehen auftreten kann. Andererseits sind bei 15 bis 20% aufgeklärter Varianz die Rahmenbedingungen und die pädagogischen Orientierungen für die Ausgestaltung der Unterrichtsqualität und für mögliche Verbesserungsmaßnahmen nicht bedeutungslos. Interpretiert man vorsichtig die Regressions-Faktor-Struktur-Koeffizienten, so zeigt sich für den Gesamtwert der Unterrichtsqualität folgendes Bild:<sup>5</sup> Der positive Pol der Vorhersagevariablen kann beschrieben werden durch Klassenlehrerinnen, die in ihren pädagogischen Orientierungen der Förderung der kindlichen Kreativität einen höheren Stellenwert einräumen und die früher als andere Lehrerinnen von den Schülern die Bewältigung spezifischer Entwicklungsaufgaben erwarten, durch vollbeschäftigte Klassenlehrerinnen (höhere Unterrichtsverpflichtungen) und durch eine bessere Ausstattung der Klassenräume mit verschiedenen Materialien. Der negative Pol der Vorhersagevariablen wird durch die umgekehrten Merkmalsausprägungen bestimmt. Variablen wie die Berufserfahrung als Lehrerin, das Fortbildungseengagement, der Umfang der Vorbereitungszeit pro Woche, die Klassengröße und der Schülerinnenanteil - diese 5 Variablen gehören ebenfalls zu dem Prädiktorensatz - stehen in keinem Zusammenhang mit den verschiedenen Skalen der Unterrichtsqualität.

#### *4.4 Beziehungen zwischen Unterrichtsqualität und Schulleistungen im Alter von 8 Jahren*

In einem blockweise hierarchischen Regressionsmodell wurde zuerst die Bedeutung der Ebene der Grundschulklasse insgesamt geschätzt (ohne Tabel-

---

<sup>5</sup> Auf die Beziehungen zu den Skalen Klassenmanagement, Schülerorientiertes sozial-emotionales Klima und Anwendungsbezug kann hier aus Platzgründen nicht weiter eingegangen werden. Bei der Interpretation des Prädiktors Mittagessen in der Schule muss beachtet werden, dass dieses Merkmal zu einem sehr großen Teil mit dem Standort der Grundschulen in den alten und neuen Bundesländern konfundiert ist.

le). Die Qualität in der Grundschulklasse wird dabei indiziert durch 10 Variablen der pädagogischen Orientierungen der Klassenlehrerinnen (Entwicklungserwartungen, Leistungsorientierung, Orientierung auf Förderung der Kreativität), der Strukturqualität (Unterrichtsstunden pro Woche, Umfang der aufgegebenen Hausaufgaben, Anzahl unterschiedlicher Materialien im Klassenraum) und der Unterrichtsqualität im Engeren (Klassenmanagement, Schülerorientiertes sozial-emotionales Klima, Direktes Unterrichten und Anwendungsbezug).<sup>6</sup> Im ersten Block des hierarchischen Regressionsmodells gehen ein: Alter und Geschlecht des Kindes sowie als Indikator für den kognitiven Entwicklungsstand im Alter von 4 Jahren der Wortschatz (gemessen über den Peabody Picture Vocabulary Test; Dunn & Dunn, 1981); Merkmale der Qualitäten des familialen Erfahrungsraumes und des Kindergartens, die die Kinder als Vierjährige erlebt haben; Merkmale der Qualität des familialen Lebensraums zur Grundschulzeit. Der Satz dieser Prädiktoren erklärt 41,7% der Varianz ( $p < .001$ ; adjustiertes  $R^2$  34,7%) der über den Woodcock Johnson Test of Achievement (Woodcock & Johnson, 1989, 1990) erfassten Schulleistungen im Alter von 8 Jahren. Folgen dann im zweiten Block die genannten Merkmale der Qualität der Grundschulklasse, so werden zusätzlich 9,1% der Varianz ( $p < .05$ ; adjustiertes  $R^2$  7,5%) in den Schulleistungen erklärt. Die insgesamt niedrigere Varianzaufklärung durch die pädagogische Qualität der Grundschulklasse ist einerseits modellbedingt, da in dem hierarchischen Ansatz die erklärenden Varianzanteile, die der Block der Grundschulklasse mit dem vorherigen Block gemeinsam hat, diesem zugeschlagen werden. Andererseits liegen 9% erklärte Varianz durchaus im Bereich des bei solchen Untersuchungen Erwartbaren (vgl. z.B. Reezigt, Guldemond & Creemers, 1999; Roßbach & Tietze, 1996; Schnabel, 1998; Helmke & Weinert, 1997b). Allzu großer (schul-) pädagogischer Optimismus ist somit unangemessen, vielmehr müssen die Auswirkungen der pädagogischen Qualität auf der Ebene einer Grundschulklasse realistisch eingeschätzt werden.

In einem zweiten Schritt wurde der spezielle Beitrag der Unterrichtsqualität für die Schulleistungen im Alter von 8 Jahren analysiert. In mehr heuristischer Perspektive enthält hierzu Tabelle 3 die einfachen Korrelationen der Skalen Klassenmanagement, Schülerorientiertes sozial-emotionales Klima, Direktes Unterrichten und Anwendungsbezug mit den Schulleistungen sowie die entsprechenden Partialkorrelationen, wenn die Effekte der folgenden Variablen aus der Schulleistung auspartialisiert werden: Alter der Schüler, Bildungsstand der Mutter, häusliches Anregungsniveau während der Kindergartenzeit (gemessen über die Home Observation for Measurement of the

<sup>6</sup> In Voranalysen wurden negative Korrelationen der Skala Unterstützung bei Stillarbeit zu den Kriterien gefunden, die vermuten lassen, dass nicht nur die Qualität der Stillarbeit, sondern auch das quantitative Ausmaß durch die Skala erfasst wurde. Aufgrund dieser konzeptuellen Unklarheit wird diese Skala bei den Analysen der Auswirkungen auf die Schulleistungen nicht berücksichtigt.

Environment von Caldwell & Bradley, 1984) und Schulfähigkeit des Kindes zu Ende der Kindergartenzeit (gemessen über den Duisburger Vorschul- und Einschulungstest DVET, Meis, 1973).

**Tabelle 3:**  
Korrelationen und Partialkorrelationen der Skalen der Unterrichtsqualität mit den Schulleistungen im Alter von 8 Jahren

	Korrelationen	Partialkorrelationen
Klassenmanagement	-.06	-.06
Schülerorientiertes sozial-emotionales Klima	-.10	.04
Direktes Unterrichten	.01	.06
Anwendungsbezug	.12*	.23**

\*  $p < .10$ ; \*\*  $p < .01$

Tabelle 3 weist letztlich nur auf einen Effekt hin (der genauso für die Subtests des Woodcock Johnson Test of Achievement gilt): Je ausgeprägter der Anwendungsbezug des beobachteten Unterrichts ist, desto höher fallen die Schulleistungen aus. Die Berücksichtigung der fächerübergreifenden Vernetzungen der behandelten Inhalte, Aktivitäten oder Lernaufgaben und ihrer Verbindungen zu Inhalten oder Kompetenzen außerhalb des Klassenzimmers steht somit klar - wenn auch eher niedrig - in Beziehung zu besseren Schulleistungen in der Grundschule. Keine Beziehungen bestehen zwischen den Schulleistungen und dem Klassenmanagement, dem schülerorientierten sozial-emotionalen Klima und dem direkten Unterrichten. Die hier gefundenen Ergebnisse sind nur teilweise mit jenen der SCHOLASTIK-Studie vergleichbar (Helmke & Weinert, 1997b). Dort zeigen sich - im Hinblick auf mit der vorliegenden Studie vergleichbaren Konzepten - zwar ebenfalls keine förderlichen Effekte des Sozialklimas, wohl aber klare, wenn auch niedrige positive Beziehungen von Aspekten des direkten Unterrichtens und der Klassenführung zu den Leistungszuwächsen in Mathematik von Beginn der 3. Klasse bis zur 4. Klasse. Keine Korrelationen wurden mit den Leistungszuwächsen in Rechtschreiben gefunden. Demgegenüber zeigen sich in der vorliegenden Untersuchung keine Auswirkungen des direkten Unterrichtens und des Klassenmanagements auf die Schulleistungen im Engeren. Der Anwendungsbezug des Unterrichts wurde in der SCHOLASTIK-Studie nicht erhoben.

## 5. Zusammenfassung und Diskussion

Der Beitrag ging von einem Verständnis von pädagogischer Qualität auf der Ebene einer Grundschulklasse aus, das 3 Qualitätsbereiche spezifiziert: Qualität der pädagogischen Orientierungen der Klassenlehrerinnen, Qualität der strukturellen Rahmenbedingungen einer Klasse und Qualität der Unterrichtsprozesse. Auf der Basis von - messtechnisch zufrieden stellenden - externen

Beobachtungen konnten faktorenanalytisch 5 Skalen der Unterrichtsqualität identifiziert werden, die weitgehend unabhängig voneinander sind: Klassenmanagement, Schülerorientiertes sozial-emotionales Klima, Direktes Unterrichten, Anwendungsbezug und Unterstützung bei Stillarbeit. Die Ergebnisse der Untersuchung in 214 Grundschulklassen weisen auf einen Verbesserungsbedarf in der Praxis hin. Dies gilt insbesondere für Verbesserungen im schülerorientierten sozial-emotionalen Klima und im Anwendungsbezug des Unterrichts, d.h. im Aufweisen von fächerübergreifenden Vernetzungen der Lernaufgaben und ihres Bezugs zu realen Anwendungsmöglichkeiten, wobei von einer Verbesserung des Anwendungsbezugs eine Steigerung der Lernmotivation der Schüler wie auch ein verbesserter Gebrauch des gelernten Wissens erwartet wird. Die Unterrichtsqualität hängt zwar mit den Rahmenbedingungen der Klasse und den pädagogischen Orientierungen der Klassenlehrerinnen zusammen, der Zusammenhang muss aber als eher niedrig eingeschätzt werden. Somit können Verbesserungen der Unterrichtsqualität nur sehr eingeschränkt über z.B. verbesserte Rahmenbedingungen des Unterrichts erreicht werden. Vielmehr scheint es erforderlich, unmittelbar am Lehrerverhalten anzusetzen, z.B. durch Trainings entsprechender Verhaltensweisen.

Auf der Ebene der untersuchten 272 Schüler konnte zwar ein Zusammenhang zwischen der Qualität einer Grundschulklasse und den Schulleistungen im Alter von 8 Jahren aufgewiesen werden, der aber - ähnlich wie in vergleichbaren Untersuchungen (vgl. z.B. Clausen, 2000; Gruehn, 2000; Kyriakides, Campbell & Gagatsis, 2000) - eher niedrig ausfällt. Im Hinblick auf die einzelnen Dimensionen der Unterrichtsqualität konnte nur ein - ebenfalls niedriger - Effekt des Anwendungsbezugs auf die Schulleistungen festgestellt werden: Je mehr Wert auf diesen Anwendungsbezug bei der Behandlung von Lernaufgaben gelegt wird, desto besser fallen die Schulleistungen aus. Erklärungsbedürftig sind die Unterschiede zur SCHOLASTIK-Studie, in der positive Effekte der Klassenführung und des direkten Unterrichtens gefunden wurden. Eine mögliche Erklärung für den fehlenden Effekt des direkten Unterrichtens könnte darin bestehen, dass in der SCHOLASTIK-Studie 3. und 4. Klassen untersucht wurden, während hier der Unterricht in der Regel in 2. Klassen beobachtet wurde. Es könnte sein, dass in der Anpassungsphase der Schüler an die Anforderungen der Grundschule ein direktes Unterrichten noch nicht so bedeutsam ist wie in den späteren Jahren der Grundschule. Insgesamt stellt sich zudem die Frage, ob die in der vorliegenden Untersuchung - wie in vielen anderen Untersuchungen - fachunspezifisch definierten allgemeinen Merkmale der Unterrichtsqualität nicht zu breit sind und ob sie nicht einer mehr fachdidaktischen Auslegung und einer auf den Unterricht in einem spezifischen Fach bezogenen Ausdifferenzierung bedürfen (vgl. Einsiedler, 2000).

Bei der Interpretation der vorliegenden Ergebnisse müssen weiterhin methodische Einschränkungen berücksichtigt werden. Die Schulleistungen der Schüler im Alter von 8 Jahren konnten nur über einen standardisierten Test erhoben werden, der möglicherweise zu unsensibel für die Auswirkungen



des Unterrichts in den deutschen Grundschulklassen ist. Aufgrund der Anlage der ECCE-Studie konnte die Unterrichtsqualität nur einmal (vgl. aber die Angaben zur Stabilität) und nicht fachspezifisch erfasst werden. Zudem wurde nur die Sicht externer Beobachter thematisiert; eine mehrperspektivische Erhebung der Unterrichtsqualität (aus Schüler- und Lehrersicht und aus der Sicht externer Beobachter), die vielleicht die Vorhersagekraft steigert, war nicht möglich (vgl. hierzu Clausen, 2000). Schließlich sind die Analysen der Auswirkungen der Unterrichtsqualität auf die Schulleistungen nur auf Individualebene möglich (meistens liegen nur Daten von einem Schüler pro Klasse vor), d.h., es können keine Effekte auf der Ebene der Klassen untersucht werden. Es gibt aber Hinweise, dass höhere Effekte auf der Klassenebene gefunden werden können (vgl. Gruehn, 2000). Vor diesem Hintergrund wird deshalb nicht davon ausgegangen, dass das Konzept der Unterrichtsqualität nur eine sehr geringe prädiktive Kraft für die Schulleistungen hat. Vielmehr sind weitere Untersuchungen erforderlich, die mit einem verfeinerten Konzept und ohne die genannten methodischen Restriktionen arbeiten.

## Literatur

- Caldwell, B.M. & Bradley, R.M. (1984). *Home Observation for Measurement of the Environment*. Little Rock, AK: University of Arkansas at Little Rock, College of Education, Center of Child Development and Education.
- Clausen, M. (2000). *Wahrnehmung von Unterricht. Übereinstimmung, Konstruktvalidität und Kriteriumsvalidität in der Forschung zur Unterrichtsqualität*. Dissertation am Fachbereich Erziehungswissenschaft und Psychologie der FU Berlin.
- Dunn, L.M. & Dunn, L.M. (1981). *Peabody Picture Vocabulary Test - Revised*. Manual for forms L and M. Circle Pines. Minnesota: American Guidance Service.
- Einsiedler, W. (1997). Unterrichtsqualität und Leistungsentwicklung: Literaturüberblick. In F.E. Weinert & A. Helmke (Hg.), *Entwicklung im Grundschulalter* (S. 225-240). Weinheim: Psychologie Verlags Union.
- Einsiedler, W. (2000). Von Erziehungs- und Unterrichtsstilen zur Unterrichtsqualität. In M.K.W. Schweer (Hg.), *Lehrer-Schüler-Interaktion* (S. 109-128). Opladen: Leske+Budrich.
- European Child Care and Education (ECCE) - Study Group (1999). *School-age assessment of child development: Long-term impact of pre-school experiences on school success, and family-school relationships*. Report submitted to the European Union DG XII: Science, Research and Development. RTD Action: Targeted Socio-Economic Research. ([http://improving-ser.sti.jrc.it/default/show.gx?Object.object\\_id=TSER----000000000000635&\\_app.page=show-TSR.html](http://improving-ser.sti.jrc.it/default/show.gx?Object.object_id=TSER----000000000000635&_app.page=show-TSR.html))
- Gerstenmaier, J. & Mandl, H. (1995). Wissenserwerb unter konstruktivistischer Perspektive. *Zeitschrift für Pädagogik*, 41, 867-888.
- Gruehn, S. (2000). *Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung*. Münster: Waxmann.
- Helmke, A. & Schrader, F.W. (1997). Unterrichtsbeurteilungen durch externe Beobachter. In F.E. Weinert & A. Helmke (Hg.), *Entwicklung im Grundschulalter* (S. 510-514). Weinheim: Psychologie Verlags Union.
- Helmke, A. & Weinert, F.E. (1997). Bedingungsfaktoren schulischer Leistungen. In F.E. Weinert (Hg.), *Psychologie des Unterrichts und der Schule*. Enzyklopädie

- der Psychologie, Serie Pädagogische Psychologie. Bd. 3 (S. 71-176). Göttingen: Hogrefe. (a)
- Helmke, A. & Weinert, F.E. (1997). Unterrichtsqualität und Leistungsentwicklung: Ergebnisse aus dem SCHOLASTIK-Projekt. In F.E. Weinert & A. Helmke (Hg.), *Entwicklung im Grundschulalter* (S. 241-251). Weinheim: Psychologie Verlags Union. (b)
- Kyriakides, L., Campbell, R.J. & Gagatsis, A. (2000). The significance of the classroom effect in primary schools: An application of Creemers' comprehensive model of educational effectiveness. *School Effectiveness and School Improvement*, 11, 501-529.
- Meis, R. (1973). *Duisburger Vorschul- und Einschulungstest (DVET)*. Weinheim: Beltz.
- Reezigt, G.J., Guldemon, H. & Creemers, B.P.M. (1999). Empirical validity for a comprehensive model of educational effectiveness. *School Effectiveness and School Improvement*, 10, 193-216.
- Roßbach, H.G. & Tietze, W. (1996). *Schullaufbahnen in der Primarstufe. Eine empirische Untersuchung zu Integration und Segregation von Grundschulern*. Münster: Waxmann.
- Sammons, P., Hillman, J. & Mortimore, P. (1995). *Key characteristics of effective schools. A review of school effectiveness research*. University of London, Institute of Education.
- Scheerens, J. (1992). *Effective schooling. Research, theory and practice*. London: Cassell.
- Schnabel, K.U. (1998). Schuleffekte. In D.H. Rost (Hg.), *Handwörterbuch Pädagogische Psychologie* (S. 431 - 435). Weinheim: Psychologie Verlags Union.
- Secada, W. G. (1997). *Instructional Environment Observation Scales*. Unveröffentl. Forschungsinstrument. University of Wisconsin-Madison.
- Tietze, W. (Hg.), Meischner, T., Gänsfuß, R., Grenner, K., Schuster, K.-M., Völkel, P. & Roßbach, H.G. (1998). *Wie gut sind unsere Kindergärten? Eine Untersuchung zur pädagogischen Qualität in deutschen Kindergärten*. Neuwied: Luchterhand.
- Wang, M.C., Haertel, G.D. & Walberg, H.J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63, 249-294.
- Weinert, F.E. (1998). Neue Unterrichtskonzepte zwischen gesellschaftlichen Notwendigkeiten, pädagogischen Visionen und psychologischen Möglichkeiten. In Bayerisches Staatsministerium für Unterricht, Kultus, Wissenschaft und Kunst (Hg.), *Wissen und Werte für die Welt von morgen. Dokumentation zum Bildungskongress des Bayerischen Staatsministeriums für Unterricht, Kultus, Wissenschaft und Kunst* (S. 101-125). München/Donauwörth: Auer.
- Woodcock, R.W. & Johnson, M.B. (1989, 1990). *Woodcock-Johnson Psycho-Educational Battery - Revised*. Allen, TX: DLM Teaching Resources.

Anschrift des Autors:

Prof. Dr. Hans-Günther Roßbach  
 Otto-Friedrich-Universität Bamberg,  
 Lehrstuhl für Elementar- und Familienpädagogik  
 Markusplatz 3, 96047 Bamberg